



**Family list****1** family member for: **JP2000215238**

Derived from 1 application

[Back to JP2000215](#)**1 METHOD FOR DETECTING ILLEGAL LITERARY WORK****Inventor:** NAKADA JUNJI**Applicant:** HITACHI LTD**EC:****IPC:** *G06F12/14*; *G06F12/00*; *G06F21/24* (+15)**Publication info:** **JP2000215238 A** - 2000-08-04

---

Data supplied from the **esp@cenet** database - Worldwide

(51) Int. Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
G 0 6 F 17/60		G 0 6 F 15/21	Z 5 B 0 1 7
12/00	5 2 0	12/00	5 2 0 E 5 B 0 4 9
12/14	3 2 0	12/14	3 2 0 E 5 B 0 8 2
H 0 4 N 1/387		H 0 4 N 1/387	5 C 0 7 6

審査請求 未請求 請求項の数7 O L (全 5 頁)

(21) 出願番号 特願平11-12732

(22) 出願日 平成11年1月21日 (1999.1.21)

(71) 出願人 000006108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 中田 順二

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(74) 代理人 100068504

弁理士 小川 勝男

最終頁に続く

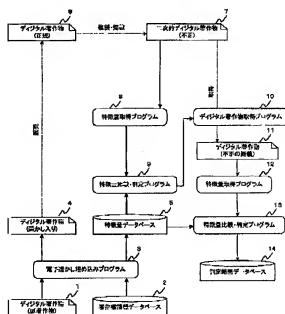
(54) 【発明の名称】 不正著作物検出方法

(57) 【要約】

【課題】電子透かし技術だけに基いた不正著作物検出方法では、インターネット上のデジタル著作物をすべて取得した上で、各々のデジタル著作物に透かし情報が含まれていないかどうかを判定する必要がある。

【解決手段】類似性が高いと判定されたデジタル著作物を取得する手段(デジタル著作物取得プログラム10)と、デジタル著作物のデータを読み取ることで初めて取得可能となる特徴量を取得する手段(特注量取得プログラム12)と、取得された特徴量と登録された特徴量とを比較して類似度を判定する第二の判定手段(特徴量比較・判定プログラム13)とを設けた。

図1



## 【特許請求の範囲】

【請求項1】複数のデジタル著作物の中から、あるデジタル著作物を元に作成された不正著作物を検出する不正著作物検出方法において、デジタル著作物固有の複数の特徴的な値を比較することにより、不正著作物であるかどうかを段階的に判定することを特徴とする不正著作物検出方法。

【請求項2】デジタル著作物固有の特徴的な値として、ファイル名称を用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項3】デジタル著作物固有の特徴的な値として、ファイルのハッシュ値を用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項4】デジタル著作物固有の特徴的な値として、ファイルのハッシュ値を用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項5】デジタル著作物固有の特徴的な値として、デジタル著作物のデータの全部、または全てを用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項6】デジタル著作物固有の特徴的な値として、デジタル著作物に埋め込まれた電子透かし情報を用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項7】デジタル著作物のデータを読み取ることなく取得可能な著作物固有の特徴的な値を先に比較し、続いてデジタル著作物のデータを読み取ることによって取得可能な著作物固有の特徴的な値を比較して不正著作物であるか否かを段階的に判定することを特徴とする請求項1記載の不正著作物検出方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明はデジタル著作物の著作権保護を目的としたもので、特に不正に作成された著作物を検出する方法に関する。

## 【0002】

【従来技術】デジタル著作物は複製や加工編集が容易なため、著作権保護技術が欠かせないものとなっている。中でも、電子透かし技術はデジタル著作物の外観にほとんど変化を与えずに情報を書き込むことが可能なため、不正著作物の検出効果が期待されている。

【0003】ところで、電子透かし技術はデジタル著作物の外観にほとんど変化を与えないため、当然ながら電子透かしが施されたデジタル著作物であるか否かを外見から判断することはできない。すなわち、電子透かし技術だけに基いて不正著作物を検出するためには、流通しているすべてのデジタル著作物に対して総当たりで電子透かしの有無をチェックする必要がある。

【0004】<http://www.mken.co.jp/dwcw.html>に記載のデジタルコンテンツ不正利用監視センターでは、この

問題に対してインターネット検索ロボットを適用することで解決を試みている。このインターネット検索ロボットはWebページ上のデジタル著作物を機械的に次々に取得し、透かし検出システムにかけることで、電子透かしを検出している。

## 【0005】

【発明が解決しようとする課題】しかし、上記従来技術では、インターネット上のデジタル著作物をすべて取得した上で、各々のデジタル著作物に透かし情報が含まれていないかどうかを判定する必要がある。これは技術的には可能であるものの、インターネット上に膨大な数のWebページが存在することを考え合わせると、現実的な手段とは言えない。

【0006】現在、インターネット接続プロバが個人用のWebページのために標準的に用意しているディスク容量は大体10MB程度である。全世界のインターネットユーザーは少なく見積もっても数千万人は存在するから、単純に計算すると、個人用のWebページだけで、数百TBの容量になる。100k/s/sの速度でWebページが取得できたとしても、数百TBのWebページを巡回するには、何百年もの時間が必要となる。これはありまにも非現実的な数値であり、最近のWebページの更新頻度から言っても、1〜2ヶ月で一通り巡回できる必要がある。

【0007】また、デジタル著作物から電子透かし情報を読み出すためには、電子透かし情報を埋め込む前の原画像が必要だったり、特殊な演算が必要だったりする。より効率的にWebページを巡回するためには、全てのデジタル著作物を取得する前や取得したデジタル著作物の電子透かしをチェックする前における程度の絞り込みを行うことが求められる。

## 【0008】

【課題を解決するための手段】以上の課題を解決するために、本発明では、著作権を保護すべきデジタル著作物の複数の特徴量を登録・管理する手段と、インターネット上に存在するデジタル著作物のデータを読み取ることなく取得可能な特徴量を取得する手段と、取得された特徴量と登録された特徴量とを比較して類似度を判定する第一の判定手段と、類似性が高いと判定されたデジタル著作物を取得する手段と、デジタル著作物のデータを読み取ることによって取得可能な特徴量を取得する手段と、取得された特徴量と登録された特徴量とを比較して類似度を判定する第二の判定手段とを設けた。

## 【0009】

【発明の実施の形態】以下、図1〜図4を用いて本発明をデジタル著作物販売に適用した場合の一実施例を説明する。図1は本発明の一実施例を構成するブロック図、図2は著作権情報データベースの例、図3は特徴量データベースの例、図4は特徴量の取得と比較・判定に関わる部分の処理フローである。

【0010】図1において、1は著作権保護の対象とするデジタル著作物の原著作物、2はデジタル著作物1に関する著作権情報を管理する著作権情報データベース、3はデジタル著作物1に対して電子透かし処理を施す電子透かし埋め込みプログラム、4は電子透かし処理の施されたデジタル著作物、5はデジタル著作物4の特徴量を管理する特徴量データベース、6はデジタル著作物4を販売した結果として購入者側に作成される正規のデジタル著作物、7はデジタル著作物6を複製することにより作成され、ネットワーク上に不正に掲載される二次的デジタル著作物、8はネットワーク上のデジタル著作物の特徴量を取得する特徴量取得プログラム、9は特徴量取得プログラム8が出力する特徴量と特徴量データベース5の内容とを比較し、類似したものがないかどうかを判定する特徴量比較・判定プログラム、10は特徴量比較・判定プログラム9の判定結果に従って呼び出されて、二次的デジタル著作物7を随時取得するデジタル著作物取得プログラム、11はデジタル著作物取得プログラム10によって取得された不正の疑義のあるデジタル著作物、12は取得したデジタル著作物11の特徴量を取得する特徴量取得プログラム、13は特徴量取得プログラム12が出力する特徴量と特徴量データベース5の内容とを比較し、該当するものがないかどうかを判定する特徴量比較・判定プログラム、14は特徴量比較・判定プログラム13の判定結果を格納する判定結果データベースである。

【0011】まず、電子透かし埋め込みプログラム3を用いて、デジタル著作物1に対して著作権情報データベース2から読み出した情報を透かし情報として埋め込む。著作権情報データベース2を構成するテーブルとしては図2に示すような3つのテーブル、すなわち著作物テーブル15、購入者テーブル16、売上テーブル17が考えられるが、透かし情報として埋め込むものは著作物テーブル15の各要素を一意に識別する著作物IDだけでもよいし、売上テーブル17の各要素を一意に識別する売上IDだけでもよい。または、著作物テーブル15に含まれる著作物名称と購入者テーブル16に含まれる購入者名称などを直接透かし情報として利用してもよい。

【0012】続いて、電子透かし埋め込みプログラム3は、透かし情報を埋め込んだデジタル著作物4を作成するとともに、デジタル著作物4に関する特徴量を特徴量データベース5に登録する。特徴量データベース5を構成するテーブルとしては図3に示すような要素をもつ特徴量テーブル18が考えられる。

【0013】特徴量テーブル18は、特徴量を一意に識別する特徴量ID、特徴量の対象となるデジタル著作物を一意に識別する著作物ID、デジタル著作物のファイル名称、ファイルサイズ、ファイルのハッシュ値、ハッシュ値を得るのに利用したハッシュ関数、ファイルの実体、透かし情報、どのような情報を透かし情報として用

いたかを表すための透かし情報の種類、から構成される。

【0014】デジタル著作物4は正規の販売手続きを経て購入者へ渡され、購入者のもとでデジタル著作物6として利用される。私的な利用の範囲にとどまる複製は特に法律で規制するものではないが、ネットワーク上へ掲載された場合は不正な二次的デジタル著作物7となる。二次的デジタル著作物7は他のデジタル著作物とともにネットワーク上に存在するため、これを識別する必要がある。この識別作業から不正に複製された著作物と判定するまでの処理フローについて図4を用いて説明する。

(S001) 特徴量取得プログラム8は、ネットワーク上のあらゆるサイトを自動的に探索するロボットプログラムで、特徴量としてファイルの名称やファイルのサイズをファイルのURLとともに取得する。

【0015】(S002) 特徴量比較・判定プログラム9は、特徴量取得プログラム8が取得したファイル名称やファイルサイズと合致するレコードが特徴量データベース5にないかどうかを検索する。ここで、ファイル名称はファイルの内容を変えずに簡単に変更できるため、ファイル名称の中でファイルの種別を表す拡張子部分のみを比較してもよい。

【0016】(S003) 特徴量データベース5に含まれるファイルサイズとファイル名称の組と同一または類似の特徴量を持つネットワーク上のデジタル著作物を、先に取得したURLからデジタル著作物取得プログラム10を用いて取得する。

【0017】(S004) 取得したデジタル著作物11のハッシュ値を、特徴量データベース5の該当するレコードに登録されているハッシュ関数を用いて算出する。

【0018】(S005) 算出したハッシュ値が、特徴量データベース5の該当するレコードと一致するかどうかを比較判定する。一致する場合は(S006)に進む。一致しない場合は処理を終了する。

【0019】(S006) デジタル著作物11と特徴量データベース5の該当するレコードに含まれるファイル実体とを比較する。一致する場合は(S007)に進む。一致しない場合は処理を終了する。

【0020】(S007) デジタル著作物11に含まれる透かし情報と特徴量データベース5の該当するレコードに含まれる透かし情報とを比較判定する。

【0021】(S008) 判定結果を判定結果データベース14に登録して処理を終了する。判定結果には、透かし情報を除いた特徴量が特徴量データベース5に登録されたものと一致するネットワーク上のデジタル著作物の所在に関する情報が格納されることになる。

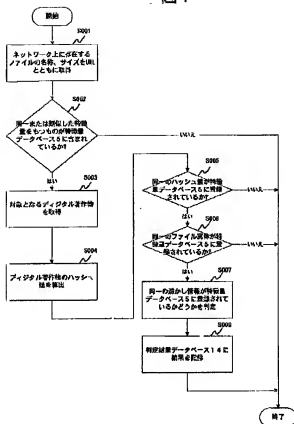
【0022】

【発明の効果】本発明によると、電子透かしを用いた不正著作物検出方法において、ネットワーク上の全てのデ



【図4】

図4



フロントページの続き

Fターム(参考) 5B017 AA06 BA05 BA07 BB03 CA16  
 5B049 AA05 BB26 CC22 CC23 DD05  
 EE03 EE05 EE56 FF09 GG04  
 GG07 GG09  
 5B082 AA13 BA07 EA09 GA02 GC03  
 GC04  
 5C076 AA14 BA06